Alternatives to Math Placement Exams: A Look at Discriminant Analysis, Neural Networks, and Ensembles of Networks

S. J. Sheel
Dept. of Computer Science
Coastal Carolina University
Conway, SC U.S.A.

R.S. Renner
Dept. of Computer Science
California State University, Chico
Chico, CA U.S.A.

S. K. Dawsey
AVX Corporation
Conway, SC U.S.A.

Abstract: *Implementing a technique that is efficient yet accurate for college student placement into the appropriate mathematics course is of significant importance. Universities often assign students to entry-level mathematics courses based on mathematics placement examination scores. In this paper, the authors examine alternative placement strategies. Using multiple regression analysis, the accumulative high school grade point average, mathematics SAT, and the final grade in Algebra II were found to be the best predictors of success on a mathematics placement examination. Using these features, entry-level mathematics placement based on neural networks is contrasted with discriminant analysis, and proposed as an alternative to testing. Results demonstrate neural networks outperform classical discriminant analysis in predicting the recommended mathematics placement. Furthermore, preliminary results suggest ensembles of networks may provide additional benefits. Consequently, a trained neural network or ensemble of networks can be an effective alternative to a written mathematics placement test.*

Keywords: neural networks, discriminant analysis, math placement testing, ensembles

1 Introduction

As technology continuously progresses, methodologies evolve to enhance our abilities to perform arduous tasks more expediently. Utilizing modern computing technologies not only makes completing tasks more efficient, but also often achieves a higher degree of accuracy than do humans. For instance, classifying students into an appropriate entry-level course is often a time-consuming method. A traditional classification scheme is distributing an exam to a student for the purpose of measuring his or her ability in a particular subject. Persons put forth much time and energy into testing a student: the development of the exam questions; the arrangement of a time and place for taking the exam; production of a physical copy of the exam; delegation of a class proctor; “grading” the exam and the actual classification of the student into the entry-level course. Several tools can be used for predictive purposes in applied research. In studies where the criterion variable is nominal rather than continuous, neural network classifiers and classical discriminant analysis can be use to predict group membership.

Currently at Coastal Carolina University, Conway, South Carolina, most incoming freshmen and transfer students are given a mathematics placement test prior to course enrollment. Students are then placed according to their performance on the placement test. The entry-level mathematics courses for placement are as follows: Math 130I, intensive College Algebra, Math 130, College Algebra, Math 131, Trigonometry and Analytic Geometry, and Math 160, Introductory Calculus. All colleges within Coastal Carolina University, except the Wall School of Business are utilizing the examination created by the mathematics department. Students entering the College of Business are assigned to a mathematics class according to their academic achievement in secondary education.

Neural networks and discriminant analysis represent techniques that can be employed for membership classification. This study compares the effectiveness of placement based on neural network classification and discriminant analysis with the more traditional manual test-based placement . The results for both the neural approach and discriminant analysis are compared to the results obtained by the mathematics placement test. If a trained neural network or prediction equation based on discriminant analysis yield statistically similar results to the mathematics placement test, then one or the other could be used for entry-level placement.

2 Related studies

Within the last decade, neural networks and classical statistical modeling have often been compared as predictors of success or failure in various disciplines. Gorr, Nagin, and Szczypula [5] conducted a comparative study of neural networks and statistical models for predicting college students' grade point averages. Discriminant analysis is included as one of the statistical models compared to a neural network. In the business realm, Ashby and Kumar [1] compare neural networks and discriminant analysis in anticipating default among high-yield bonds. Both studies conclude that the neural network serves as a better predictor than discriminant analysis. On the other hand, Wilson and Hardgrave [14] found that the discriminant analysis approach predicts graduate student success in a master's level business administration (MBA) program better than the neural network approach.

The use of neural networks as a predictor has increased over the past few years. Cripps [4] used a neural network to predict grade point averages of Middle Tennessee State University students. According to Carbone and Piras [3], neural networks are instrumental in predicting high school dropouts. In Nelson and Henriksen's study [13], a neural network uses input from student responses on a mathematics placement examination given to incoming students at Ball State University and outputs the mathematics course in which each student should be placed. The implementation of neural networks as a prediction tool for educational placement and assessment continues to increase.

3 Methodology

When using high school performance as a method for assigning college students to an entry-level mathematics course, certain performance measures are considered. In this study, grade point average (GPA), class rank, and scores reflecting Scholastic Aptitude Test (SAT), Algebra I, Geometry, advanced mathematics, and Algebra II are high school academic performance indicators that have the potential of affecting mathematics course placement. Multiple correlation coefficients adjusted to beta weights are used to analyze the variance of these predictors of scores on the mathematics placement test [6]. The three factors chosen in this study to have the most influence on mathematics placement test scores are high school GPA, SAT mathematics score, and final grade in high school algebra II.

The scale used for high school GPA and the final grade in high school algebra II is 0.0 to 4.0, with 0.0 representing the lowest possible score or an "F" and 4.0 representing a perfect score or an "A". SAT mathematics score serves as a measurement of a student's overall mathematical ability. The scale for the SAT mathematics score ranges from a low of 200 to a high of 800.

3.1 Classical statistical modeling method

According to Ashby and Kumar [1], discriminant analysis is "a statistical technique used to classify objects into distinct groups based on a set of criteria or characteristics of the objects." Fisher's Linear Discriminant Analysis (FLDA) is a favorable classification rule that works well in situations where the groups to be discriminated are linearly separable. Objects are grouped according to their "discriminant score," which is computed based on the object's observed values of the discriminating criteria or characteristics [1]. The question remains as to whether the course classifications are of a linear nature.

3.2 Neural network method

In this series of experiments, supervised feed-forward neural networks were chosen as the network model. The networks are trained using data for entry-level college students grouped according to their performance on the mathematics placement exam. The student's high school grade point average, SAT mathematics score, and high school algebra II score are used in the input layer. The output layer consists of four nodes representing courses in which the student may be placed into by the mathematics placement exam. In this experiment, recurrent back-propagation is used to train the network. BrainMaker Professional, California Scientific Software, is used to create the neural network [2].

3.3 Ensemble method

A neural network ensemble is a collection of networks, referred to as 'students'. Ensembles have demonstrated significant improvement over single networks for classification [7], regression [8], and overall performance in terms of generalization and stability [9,12]. Ensembles may be composed of weak or well-trained students, wherein each student may be independently linear or non-linear. In such a system, when presented with identical input vectors, each student will derive a unique prediction (see figure 1).

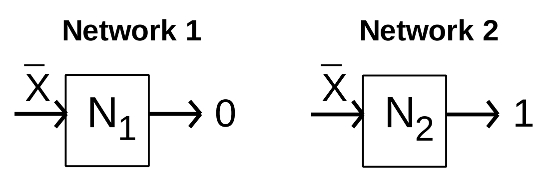

Figure 1 - Individual Network

These predictions are combined using hard or soft decision rules to form a global prediction. Figure 2 demonstrates an ideal ensemble for a sample 2-dimensional input space and associated output classification marked by Xs and Os. Each student in the ensemble is linear and is represented by *Network 1*, *Network 2*, or *Network 3*. Note that each student network has successfully landed in a unique region of the feature space. Simple voting easily resolves the problem and correctly classifies all patterns in the sample space. Key to the success of such a system is the uniqueness that each student brings to the ensemble. Empirical and theoretical evidence has shown that ensembles exhibiting a high degree of diversity or 'ambiguity' among students lead to well-trained non-linear systems and improved error-reduction [12].

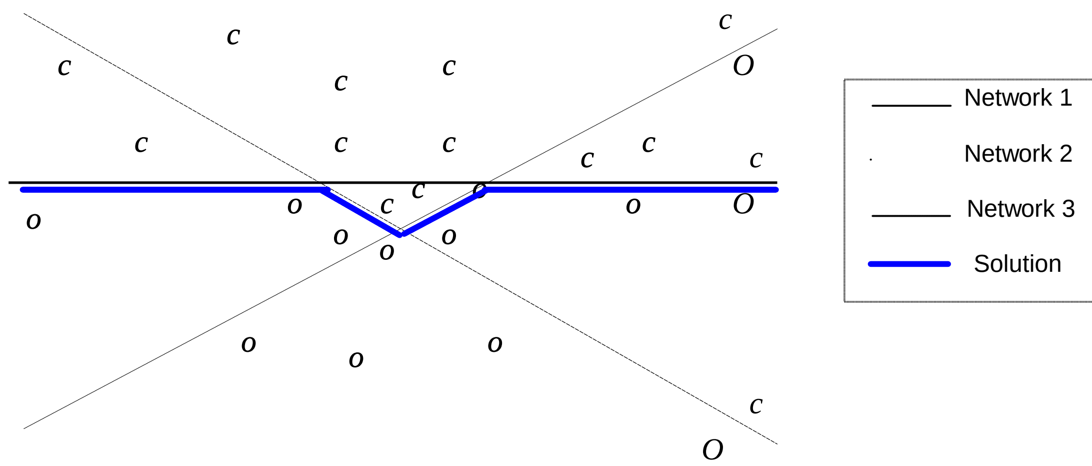

Figure 2 - linear classifiers, converge by voting

Based on previous evidence and experience with ensemble classification [9,11], it is expected that further improvement can be achieved beyond those of successful individual networks, by the replication of these networks for the generation of ensembles.

3.4 Data collection

Academic information for students entering Coastal from the fall semester of 1995 to the spring semester of 1997 was obtained through the Office of Institutional Advancement. The data is divided into the following categories for each student: hig h school grade point average, high school rank, SAT verbal, SAT mathematics, algebra I, algebra II, geometry, advanced mathematics and mathematics placement examination scores. All student records with blank fields have been deleted from the study. The intent of this initial study is to examine students who have been correctly placed in an entry-level mathematics class, therefore, only records of students receiving a final grade of "C" or higher have been selected for the study. Consequently, this study utilizes the high school GPA, the SAT mathematics score, the final grade in algebra II and the in-house mathematics placement examination score as input, with the resulting placement as labeled output. The total number of records in the final data set is 458.

4 Results

The 458 student records were randomly ordered and assigned to two equal data sets. The first group is used to train the neural networks and to create the discriminant analysis predictive equation. The second data set is used to test the tra ined neural networks and the predictive discriminant analysis equation. Both methodologies used the overall high school GPA, the SAT mathematics score, and the final grade in Algebra II as inputs, with the mathematics placement result as the output varia ble.

The topology chosen for the backprop network is a 3-10-1 network, with ten nodes in the hidden layer. The recurrent back-propagation network correctly placed 206 out of the 229 (89.9%) exemplars in the test set. The predictive equations derived using discriminant analysis correctly placed 155 out of 229 (67.7%) exemplars in the test set.

The classification rates reflect a significant difference between the neural network and discriminant analysis approach, at a 22.2% improvement using neural networks. These results suggest a dependency upon method, between the neural network and the discriminant analysis approach. These perceptions are reinforced by a chi-square contingency test, which is used to reject the null hypothesis of independence between the placement results and the type of methodology used. The placement results are dependent on methodology at the .01 level. The difference between the neural network classification results and the placement test assessment are less significant, however, which serves to validate the neural network approach as a method of choice.

The placement data used in this study proved to be significantly non-linear, as indicated by the discriminant analysis predictive equation and the prediction accuracy for *linear classifiers*, which fell to 53.67% accuracy. However, when ten linear classifiers were combined to form an ensemble of students, the course classification rate jumped to 85.26%. This ensemble classifier was generated using the Neural Network Ensemble Simulator (NNES) [10], for rapid generation of *weak classifiers* using the Adeline learning algorithm. Preliminary results provide support for further investigation of this data using NNES ensembles *of well-trained classifiers,* where results for similarly complex classification data have demonstrated a 24.47% improvement over ensembles of weak classifiers generated under identical conditions [11].

5 Conclusion

The recurrent backprop neural network significantly outperformed the predicative discriminant analysis equation as a tool for placing incoming freshman into entry-level mathematics courses. This result comes as expected, given that the FFNN has the potential for modeling non-linearities.

Generating a trained network for a particular problem domain is not always an easy task. It requires trial and error configuration attempts to settle on a network architecture and parameters sufficient for training, as in the case of the 3-10-1 network used here. Furthermore, the degree of non-linearity of the problem contributes to the complexity and training time of the overall network or system of networks. Such architectural dilemmas and time constraints can be addressed by employing constructive networks or an ensemble of weak classifiers.

When contrasted to the logistical difficulties of testing, scoring, and reporting the scores on a mathematics placement test, the neural network approach is more efficient and 89.9% effective. The empirical evidence presented here suggests that well-trained neural networks or ensembles of classifiers can be employed as an alternative to testing, for placement of students in entry-level mathematics courses.

References

[1] D. Ashby and N. Kumar. "A comparison of neural networks and classical discriminant analysis in anticipating default among high-yield bonds." *Association for Information Systems 1996 Americas Conference*

[On-line],1-4, August 1996. http://hsb.baylor.edu/ramsower/ais.ac.96/papers/ashby.htm

[2] *BrainMaker Professional User's Guide and Reference Manual*, 4th Ed., California Scientific Software, Nevada City, CA, 1993.

[3] V. Carbone and G. Piras. "Palomar project: Predicting school renouncing dropouts, using the artificial neural networks as a support for educational policy decisions." *Substance Use & Misuse*, 33 (3), 717-750, 1998.

[4] A. Cripps. "Predicting grade point average using clustering and ANNs." *Proceedings of World Congress on Neural Networks*, 486-490, 1996.

[5] W.L. Gorr, D. Nagin, and J. Szczypula. "Comparative study of artificial neural network and statistical models for predicting student grade point averages." *International Journal of Forecasting*, *10* (2), 17-34, 1994.

[6] W.H. Guertin and J.P. Bailey. *Introduction to Modern FACTER ANALYSIS*, Edward Brothers, Inc., Ann Arbor, MI, 1970.

[7] L.K. Hansen and P. Salamon, "Neural network ensembles", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp.993-1001, 1990.

[8] M.P. Perrone and L.N. Cooper, "When networks disagree: Ensemble method for neural networks," in *Artificial Neural Networks for Speech and Vision, Neural Computing Series*, chapter10, R.J. Mammone, Ed.: Chapman-Hall, 1993,.

[9] R.S. Renner. *Improving Generalization of Constructive Neural Networks Using Ensembles*, dissertation in Compu ter Science, Tallahassee: Florida State University, 1999, pp. 248.

[10] R.S. Renner, R.C. Lacher, B.J. Juliano. "A Simulation Tool for Managing Intelligent Ensembles," *Proceedings of the International Conference on Artificial Intelligence*, 578-584, 1999.

[11] R.S. Renner and R.C. Lacher. "Combining Constructive Neural Networks for Ensemble Classification," *Proceedings of the Fifth Joint Conference on Information Sciences*, 887-891, 2000.

[12] P. Sollich and A. Krogh. "Learning with ensembles: How over-fitting can be useful," in *Advances in Neural Information Processing Systems 8*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds. Cambridge, MA: MIT Press, 1996.

[13] C. Van Nelson and L.W. Henriksen. "Using neural net technology to enhance the efficiency of a computer adaptive testing application." Mid-Western Educational Research Association Annual Meeting, 1-40, 1994.

[14] R.L. Wilson and B.C. Hardgrave. "Predicting graduate student success in an MBA program: Regression versus classification." *Educational and Psychological Measurement, 55* (2), 186-195, 1995.